



Teacher and Principal Observations

Concurrent Session 3-A

Tony Milanowski
Wisconsin Center for Education Research
University of Wisconsin – Madison
&
Patrick Schuermann
Vanderbilt University

Teacher Incentive Fund Grantee Meeting
June 5 & 6, 2008
Bethesda, MD

In this age of value-added measurement, why bother with observing teachers and principals?

- Observations take time
- Require observer training
- Subject to considerable leniency bias
- Require some idea of what to look for → What is good instruction?
- Past research on teacher evaluation painted a bleak picture of reliability & validity

On the other hand...

- Not all teachers teach tested subjects or grades
- The behavioral aspects of teaching “the how” can be as important as the outcomes “the what”
- “Line of sight” shorter to teachers’ own behavior
- Teachers need formative feedback on what they need to do (behaviors, task strategies) to improve student achievement

A Few Promising Developments:

- Research on NBPTS certification
- Research on evaluation systems conducted by the Consortium of Policy Research in Education (CPRE)
- Teacher Advancement Program (TAP)
- Measurement based on “theories of instruction” (CLASS & SII)

Key: Rubrics must specify levels of performance with enough detail to make it clear what behaviors are required to be considered a great performer.

Good news: Standards-based evaluation scores can be positively related to measures of value-added student achievement (Milanowski, 2004).

Efforts to provide valid reliable measures of principal performance:

- **The ISLLC Standards (1996)**
 - 184 elements defining the knowledge, skills, and dispositions of school leaders within 6 key domains.
- **Hessel and Halloway (2002)**
 - A framework that includes rubrics for the 6 ISLLC standards and ties them together through 4 recurrent themes. For each standard and related theme, there are 4 levels of performance.
- **Reeves (2004)**
 - Includes 10 dimensions of leadership performance and provides rubrics to assess 4 levels of principal performance.
- **Vanderbilt Assessment of Leadership in Education (VALEd)**
 - Assesses leadership behaviors defined by the intersection of 6 core components of school performance and 6 key processes. 360 degree feedback tool that requires data sources to be provided.

What are some of the major issues associated with observational measures?

- Do they measure the “right stuff”?
- Formative vs. summative tension
- “Good enough” to be used for pay?
- How can they best complement value-added and other outcome measures?
- Are they useful at all for teachers and principals?

Some Features of Current Best Practice:

- Instrumentation structure and content
 - “Standards”
 - Multi-level behavioral rating scales
 - Some attempt to look at content & content-related pedagogy
- Additional forms of evidence (artifacts)
- Differentiation: processes for new, experienced, and struggling teachers

More Features of Current Best Practice:

- Frequency of classroom observation
 - Peak vs. typical performance
 - ✓ Once is not enough... nor is twice
 - Time of year & time of day effects
- Number and type of persons serving as evaluators
 - Multiple (2) observers, though not necessarily at any one time
 - Administrators, peers, district specialists, & other outsiders
 - Observer motivation & accountability
- Extent and content of evaluator training

Observational Measures Used in TIF Sites

System	Round 1 Grantees	Round 2 Grantees
NIET/TAP	5	3
State Systems	5	2
Locally Developed	4	8
Framework for Teaching	1	1
Other	1	-
Unspecified/ Undecided	2	1

Features of 9 Teacher Evaluation Systems From Round 1 TIF Sites

Characteristic	Number of Systems
Multi-level behavioral rating scale	6 (3-7 levels)
Number of observations	1 to 4; mode =1
Data collection includes artifacts	5 require, 4 permit
Multiple observers	2
Different process based on teacher experience or performance	6
Used for pay	2 systems (Denver, TAP)

What can we do to reduce the overhead?

- Focus instrumentation on key performance dimensions
 - that most affect student achievement
 - that align with district goals and represent desired district instructional strategies
- Observing classroom performance
 - More frequent shorter vs. fewer longer?
 - Tie to standards-based instructional units
- Digitize and rate artifacts & videos over the summer

Importance of Implementation Quality

- Timelines, requirements, & forms
- Evaluator training and retraining
- Evaluation motivation & accountability
- Coaching & feedback
- Teacher orientation
- Minimize teacher burden
- Support from the top
- HR system alignment

For more information

- <http://cecr.ed.gov/guides/teacherObservation.cfm>
(CECR resource on teacher observation)
- <http://cecr.ed.gov/guides/principalCompensation.cfm>
(CECR resource on compensating principals)
- http://cpre.wceruw.org/papers/HR_alignment.pdf
- Heneman, H.G. III, and Milanowski, A.T. (2004) Alignment of human resource practices and teacher performance competency. *Peabody Journal of Education*, 79:4, 108-125.